42

i	1	2	3	4	5	6	7	8	9
$y^{(i)}$	0	0	0	0	0	1	1	1	1
$g(x^{(i)})$	0.75	0.15	0.11	0.23	0.09	0.1	0.66	0.82	0.5

$$f(x) = I(g(x) \geq 0.5) = \begin{cases} 1, & g(x) \geq 0.5 \\ 0, & \text{otherwise} \end{cases}$$

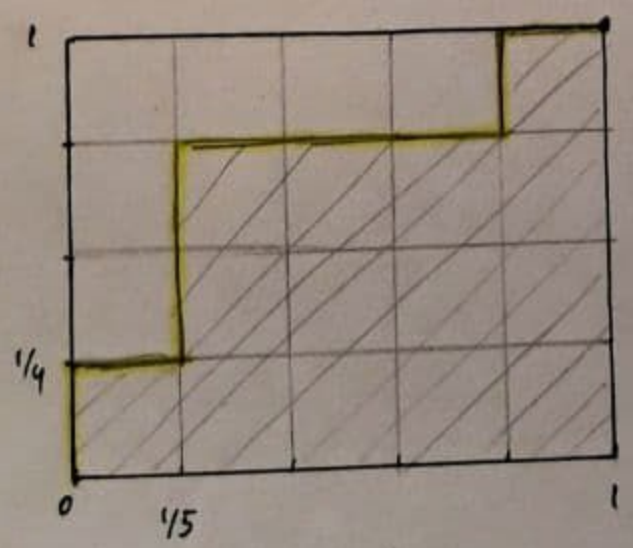
g \ y	0	1
0	TN: 4	FN: 1
1	FP: 1	TP: 3

- $FPR = \frac{FP}{TN+FP} = \frac{1}{5}$
- $FNR = \frac{FN}{FN+TP} = \frac{1}{4}$
- $TNR = 1 - FPR = 4/5$
- $TPR = 1 - FNR = \frac{3}{4}$ (recall)
- $PPV = \frac{TP}{FP+TP} = \frac{3}{4}$ (precision)
- $accuracy = \frac{TP+TN}{P+N} = \frac{7}{9}$
- $error = 1 - accuracy = \frac{2}{9}$
- $F1 = \frac{2 \cdot PPV \cdot TPR}{PPV + TPR} = \frac{2 \cdot 3/4 \cdot 3/4}{3/4 + 3/4} = \frac{3}{4}$

$N = 5$
$P = 4$

i	$\hat{y}$	y
5	0.09	0
6	0.1	1
3	0.11	0
2	0.15	0
4	0.23	0
9	0.5	1
7	0.66	1
1	0.75	0
8	0.82	1

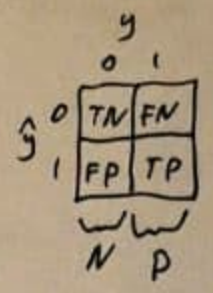
Сорт. по возраст



- $AUC = \frac{14}{20} = 0.7$

41

1) PPV_1 (precision) $= PPV_2$
TPR_1 (recall) $= TPR_2$ $\Rightarrow$ $TNR_1 = TNR_2$
$NPV_1 = NPV_2$



На одной выборке $\Rightarrow$ N, P - знаем

$$PPV = \frac{TP}{FP+TP} \Rightarrow \begin{matrix} \cdot PPV_1 = PPV_2 \\ \cdot TP_1 = TP_2 \end{matrix} \Rightarrow FP_1 = FP_2, \quad TN+FP = N \Rightarrow TN_1 = TN_2$$

$$TPR = \frac{TP}{FN+TP} = \frac{TP}{P} \Rightarrow TPR_1 = TPR_2 \Leftrightarrow TP_1 = TP_2$$
$$P = FN+TP$$
$$\Rightarrow FN_1 = FN_2$$

$$TNR = \frac{TN}{N} \text{ тогда } TN_1 = TN_2 \Rightarrow TNR_1 = TNR_2$$

$$NPV = \frac{TN}{TN+FN} \text{ тогда } \begin{matrix} TN_1 = TN_2 \\ FN_1 = FN_2 \end{matrix} \Rightarrow NPV_1 = NPV_2$$

■

2) $TNR_1 = TNR_2$
$NPV_1 = NPV_2$ $\Rightarrow$ $PPV_1 = PPV_2$
$TPR_1 = TPR_2$

Аналогично 1): $TNR_1 = TNR_2 \Rightarrow TN_1 = TN_2 \Rightarrow FP_1 = FP_2$
и
$NPV_1 = NPV_2 \Rightarrow FN_1 = FN_2 \Rightarrow TP_1 = TP_2$

Тогда $PPV_1 = PPV_2$ и $TPR_1 = TPR_2$ ■

3) Совпадение RoC кривых $\Rightarrow$ совпадение precision / recall кривой

└ FPR / TPR кривая

- $TPR = \frac{TP}{P}$
- $FPR = \frac{FP}{N}$

- $PPV(\text{precision}) = \frac{TP}{TP+FP}$
- $TPR(\text{recall}) = \frac{TP}{P}$

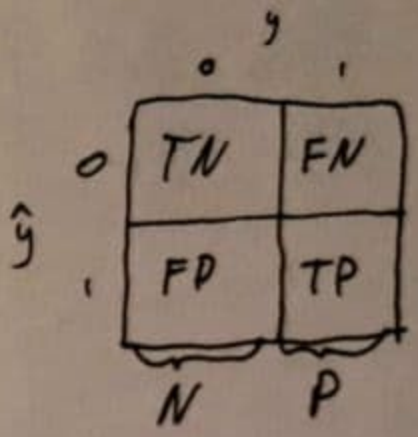
совпадение ROC $\Rightarrow TP_1 = TP_2, FP_1 = FP_2 \Rightarrow$ совпадение PPV / TPR прямой

■

40

По любым из 2-х показателей определяются остальные 2 с известными N и P

- TPR
- TNR
- PPV
- NPV



$TN+FP=N$

$FN+TP=P$

если знаем: то знаем и

- $TPR=\frac{TP}{P}$ → TP, FN
- $TNR=\frac{TN}{N}$ → TN, FP
- $PPV=\frac{TP}{TP+FP}$
- $NPV=\frac{TN}{TN+FN}$

- Легко заметить, что по TPR и TNR восстанавливаются PPV.

- По TPR и NPV:

$$\frac{TN}{TN+FN}\,a \Rightarrow TN=\frac{aFN}{(1-a)},\ FP=N-TN \rightarrow \text{восстанавливаются TNR и PPV}$$

- Аналогично по TNR и PPV восстанавливаются TRR и NPV

- По PPV и NPV:

$$\frac{TP}{TP+FP}=a \Rightarrow TP=\frac{aFP}{1-a} \rightarrow FN=P-TP$$

$$\frac{TN}{TN+FN}=b \Rightarrow TN=\frac{bFN}{1-b} \rightarrow FP=N-TN$$

$$\rightarrow \text{восстанавливаются TPR и TNR}$$

Артамонов

15

X_1	0	0	1	1	0	0	1	1	1	0
X_2	0	1	0	1	1	1	1	1	1	1
y	0	0	0	0	0	1	1	1	1	1

$y=0$	0.5
$y=1$	0.5

$$P(y=0 \mid X_1=1, X_2=1) = \frac{P(y=0)\,P(X_1=1, X_2=1 \mid y=0)}{P(X_1=1, X_2=1)} = \frac{P(y=0)\,P(X_1=1 \mid y=0)\,P(X_2=1 \mid y=0)}{P(X_1=1, X_2=1 \mid y=0)\,P(y=0) + P(X_1=1, X_2=1 \mid y=1)\,P(y=1)}$$

(1/2, 2/5, 3/5)

$$= \frac{6/50}{\frac{3}{10} + 6/50} = \frac{6}{21}$$

(1/2, 3/5, 1)

$$P(y=1 \mid X_1=1, X_2=1) = \frac{P(y=1)\,P(X_1=1 \mid y=1)\,P(X_2=1 \mid y=1)}{P(X_1=1, X_2=1)} = \frac{\frac{3}{10}}{\frac{3}{10} + \frac{6}{50}} = \frac{15}{21}$$